# Derivation of Ionosphere Constants
# And
# Reconstruction of Impulsive Signals
# From
# FORTE Satellite Data


*William C. Myre*
*Sandia National Laboratory (retired)*

# Introduction

It is well known (e.g., Budden, 1985, and Davies, 1965) that the ionosphere introduces frequency-dependant delays for radio signals that transit the ionosphere. For frequencies close to the plasma frequency these delays can become substantial. The difference in delays of the Ordinary (O) and Extra Ordinary (X) signals can also become large. For these frequencies, an additional delay that is a function of an inverse fourth power of frequency can be observed. This additional delay is a function of the peak electron density (Roussel-Dupre et al., 2001). Thus, by measuring the delay as a function of frequency over a large range of frequencies above the plasma frequency it is possible to estimate the slant total electron content (tec), gyro frequency (g) and peak electron density (Ne) of the ionosphere.

The FORTE satellite (Massey et al., 1998, Jacobson et al., 1999) is an excellent platform to make measurements of these delays. Previous studies (Massey et al., 1998, Roussel-Dupre et al., 2001) used short impulsive signals from the Los Alamos Portable Pulser (LAPP) (Massey et al., 1998) to derive the parameters of the ionosphere from a series of LAPP pulses as the as the satellite passed overhead.

This paper will explore the derivation of accurate parameters using the correlation of the received signal with an impulse that has transited an assumed ionosphere. With this methodology, in addition to deriving the parameters of the ionosphere, it will be shown that the pulse widths of short over-water CG- lightning pulses previously reported (Weidman and Krider, 1980, Levine et al., 1989, Jacobson, 2000 and Jacobson et al., 2001) can be determined.

# Dispersion Equation - Original Version

The basic dispersion equation derived by Roussel-Dupre et al. (2001) with different term definitions and the inverse fourth power of frequency term modified and simplified is:

$$Delay = \frac{a * tec}{f^2} \mp \frac{2 * a * g * tec}{f^3} + \frac{a * d * tec^3}{f^4}$$

Where:

Delay is in $\mu$sec

a is a constant = 13445.

tec is the slant total electron count in units of $10^{17}$ electrons / meter $^2$

f is the frequency of interest in MHz.

g is in MHz where

$g = B * \cos\Theta$

$B$ is the angular gyro frequency of an electron

$\Theta$ is the angle between the wave normal and the geomagnetic field

The double sign accounts for the ordinary(O) and extraordinary(X) waves

The equation used for d is:

$$d = \frac{60}{tec0} * \left(\frac{1}{T} - \frac{\sin(z')^2}{h}\right)$$

Where:

$60 \sim \frac{3}{2} * a * c * 10^{-5}$ (so that distances can be in $10^5$ meters)

Where: c = 300 meters / $\mu$ sec

T is the vertical slab thickness in $10^5$ meters

h is the satellite altitude in $10^5$ meters (~ 8)

tec0 is the vertical total electron count

= tec * cos (z') where:

z is the zenith angle at the event

z' is the angle between the event satellite line at the ionosphere center line and the vertical to earth.

A typical value for the ionosphere centerline is 300 km.

Since neither tec0 nor T will vary significantly during a pass (while both tec and z will), the value of d determined for one event on a pass can be used as a first estimate for the other events on the same pass. Typical values of T are from 2 to 4.

The maximum electron density (Ne) can be determined from:

$$\text{Ne} = \frac{tec0}{T} \quad \left(\text{in units of } \frac{electrons}{meter^3} * 10^{12}\right)$$

## Dispersion Equation – Assumptions Used

1) The ionosphere can be represented by a simple slab of electrons with a peak value equal to the maximum electron density (Ne) and the thickness (T) set so that the total number of electrons is the total electron count (tec).

2) The delay of an impulsive signal as a result of transiting the ionosphere can be accurately described by an equation consisting of three inverse frequency dependent terms, a quadratic, a cubic and a quartic.

3) The correct values for the coefficients for this equation are the same for the ordinary (O) and extraordinary (X) signals except the sign of the cubic term is reversed for the X signal.

4) The correct values for the coefficients can be determined by curve fitting (correlating) the signal with a signal generated by passing an impulse through an ionosphere described by the dispersion equation.

5) More accurate values for the coefficients can be found by using data from both the Low band (26 – 48 MHz)and High band (118 – 140 MHz) signals.

6) There are no significant frequency dependent delays in the FORTE Data Acquisition System

# Dispersion Equation – Required Modifications

## Slab Model

In the dispersion equation noted above the term d is not accurate since the quartic term is a function of the square of the electron density and thus using a constant term for electron density integrated across the slab thickness will cause erroneous results. As a simple modification to minimize this problem the original slab model is replaced with a triangular shaped model as shown in Figure 1. It is not necessary for the triangle to be symmetrical. With this model the quartic delay would be 2 / 3 of the value found using the original slab thickness model.

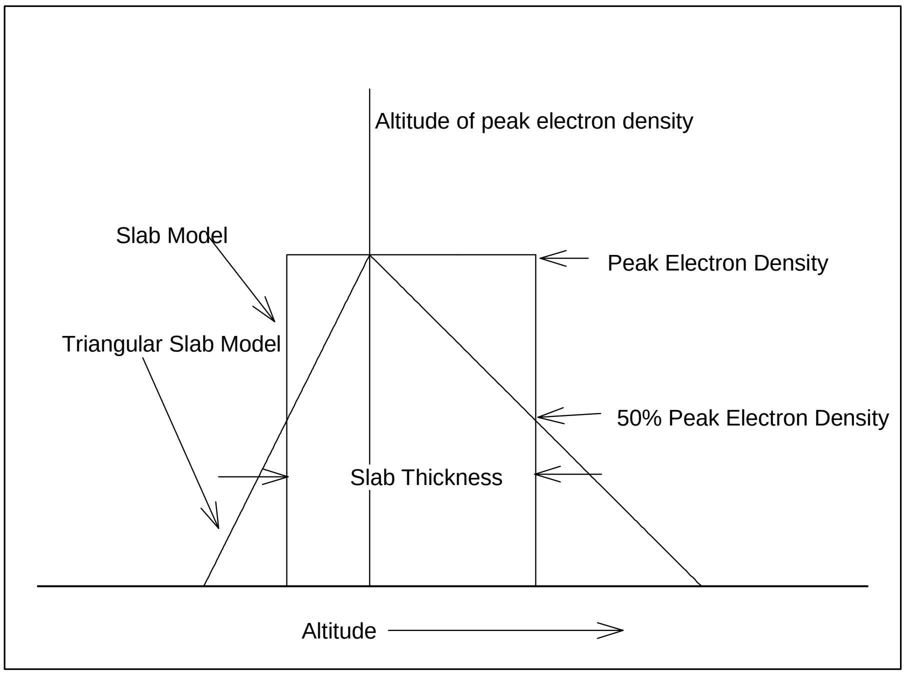


Figure 1

Roussel-Dupre et al. (2001) took the electron density data from the International Reference Ionosphere (IRI) code and then calculated values of tec, g, T, Ne, and quartic delay at 100 MHz (q) using the LANL TIPC code. Since the computations consisted of an integration of electron density along the slant path from the LAPP site to the satellite using an accurate dispersion equation, it is expected that the calculated values should be accurate.

From the derived value of q, d may be computed from

$$\text{d (actual)} = \frac{100Mhz^4 * q}{a * tec^3}$$

Then using the equation derived above with the computed values of tec and T and the zenith angle z' an estimated value of d can be determined. The ratio of these two values is plotted in Figure 2. This ratio can be used as a correction factor to arrive at a more accurate equation for d.

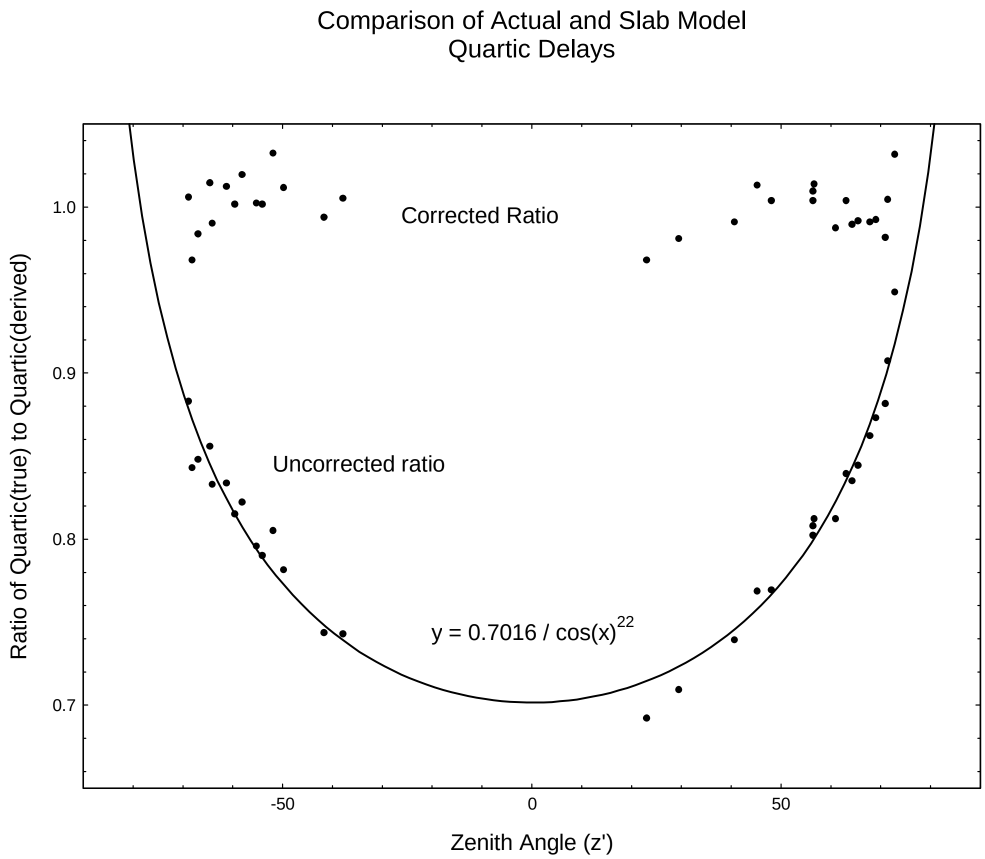


Figure 2

d (corrected): = d * k1 / k2

Where:

$k1 = 0.7016$ and

$k2 = \cos(z')^{0.22}$

While this empirically derived correction term was obtained from only 31 events on three days and near LANL, the fact that the fit is so good even though the values of tec, g and zenith angle vary widely is encouraging. Also the factor k1 is very close to 2 / 3 as expected from the triangular dispersion model.

Thus the final expression for d to be used to calculate the peak electron density and slab thickness is:

$$d = \frac{60}{tec0} * \frac{0.7016}{\cos(z')^{0.22}} * \left(\frac{1}{T} - \frac{\sin(z')^2}{h}\right)$$

## Frequency Dependant Delays

When ionosphere constants for TATR - LAPP test data were derived using only low band data, the resulting values of Ne for small tecs and zenith angles were much too large to be credible. However when data from the high band was also used, while the derived values for Ne now looked reasonable, the maximum correlation values were significantly smaller (~ 10 to 20 %). Clearly something was wrong.

The sixth assumption was changed to allow a frequency dependent delay as plotted in Figure 3. With this assumption, not only were reasonable values of Ne found at the maximum correlation, the maximum correlation was even larger than the previous maximum correlation (~ 10 – 20%).

While the cause and location of the delay (probably in more than one place) are not known, the fact that the data look so good implies they must exist.

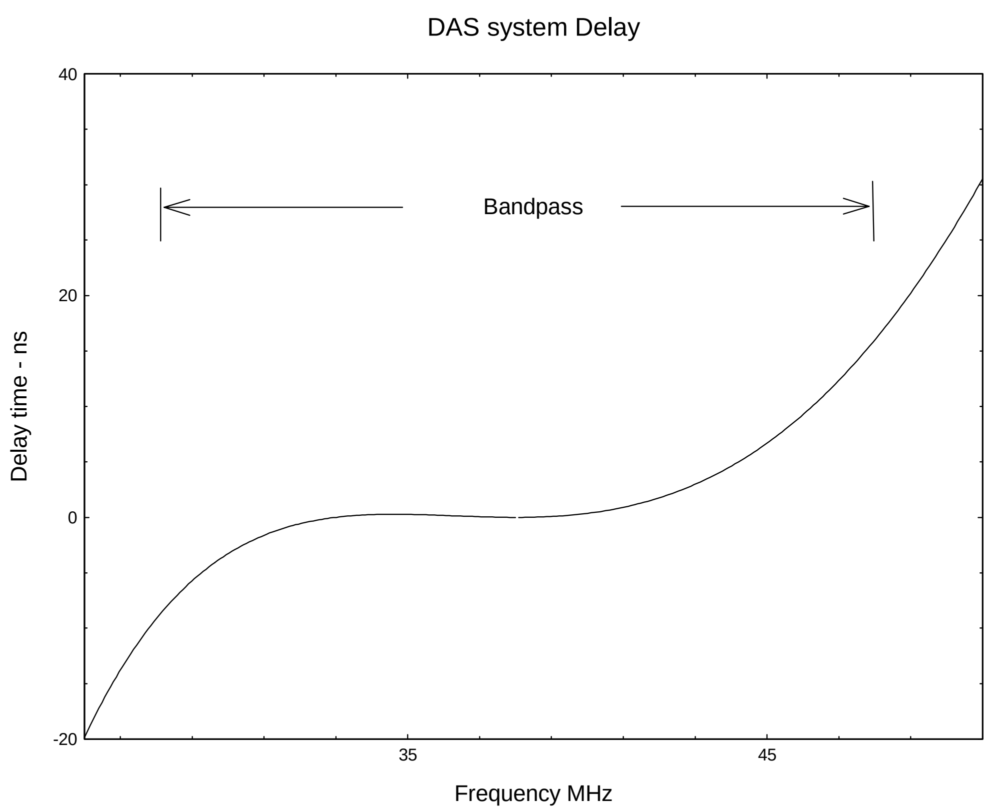


Figure 3

## Correction factor xk0

The algorithms to find the correct values for tec, g and d use the geometric mean of the maximum O and X correlation values. In deriving ionosphere constants for the 59 LAPP tests, it was found that the values of tec for the O and X maximum correlation were often slightly different than the tec from the geometric mean. This difference in the best O and X tecs varied smoothly across a pass from a positive to a negative value. Figure 4 shows this effect for a series of events on May 14. The values of the tec errors, the empirically derived value for xk0 and the values of g are plotted as a function of latitude (latitude was selected since it varies smoothly across a pass) in Figure 5.

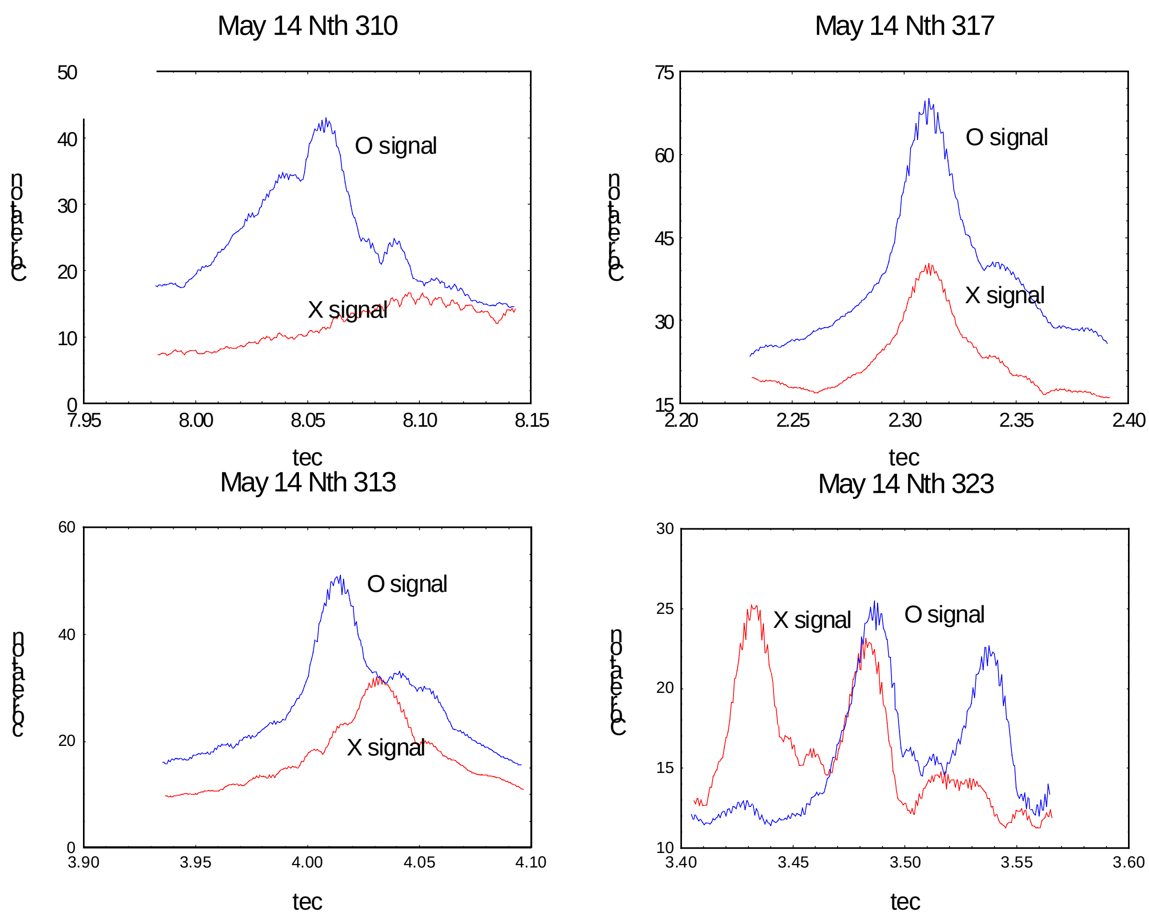


Figure 4

The tec error could be corrected by including a small $\mp 1/f^4$ term with an empirically selected coefficient. With this addition, the derived values for tec and d did not change significantly, however the value for g was decreased for events at large zenith angles. However it was noticed that the value of g derived without the inclusion of this correction were significantly larger than the correct values as found from the IRI data in the report by Roussel-Dupre et al. (2001) , (See Figure

6). It appears that this correction is proper, except that the corrected value of g is now too small for large values of tec. It seems that the use of this xk0 correction is an improvement but is obviously not the only or final correction needed. However to try to do better is probably not worth the effort. For the best estimate of g it is suggested that;

1) For values of g greater than 0.5 take the average of the values found using xk0 = 0.0 and the correct value of xk0.
2) For values of g less than 0.5 use the value found with xk0 = 0.0

Thus the required correction is:

$$Delay = \mp \frac{3 * a * xk0}{f^4}$$

A first guess for xk0 can be found from:

xk0 ~ tec * (tec * g – 1.4)

It should be noted that the xk0 correction is small. If all that is desired is a reconstruction of the original signal, it is not necessary to find xk0, but simply set xk0 to 0.0 (or a guess) and use the tec values that result in the largest correlation value.

Undoubtedly certain that this correction term is not absolutely correct. The result from including some of the terms left out in the derivation of the dispersion equation would surely be different. However it is likely that the results obtained if the correct terms were used would be similar to what is found here.

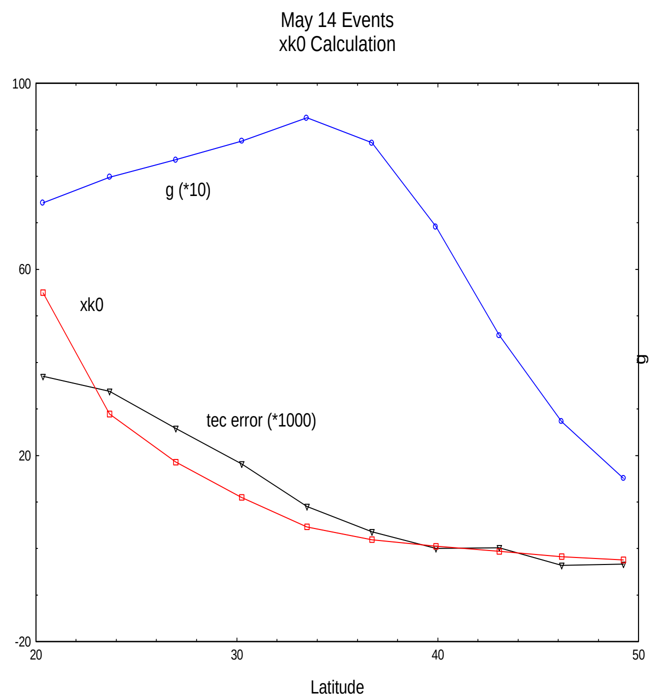


Figure 5

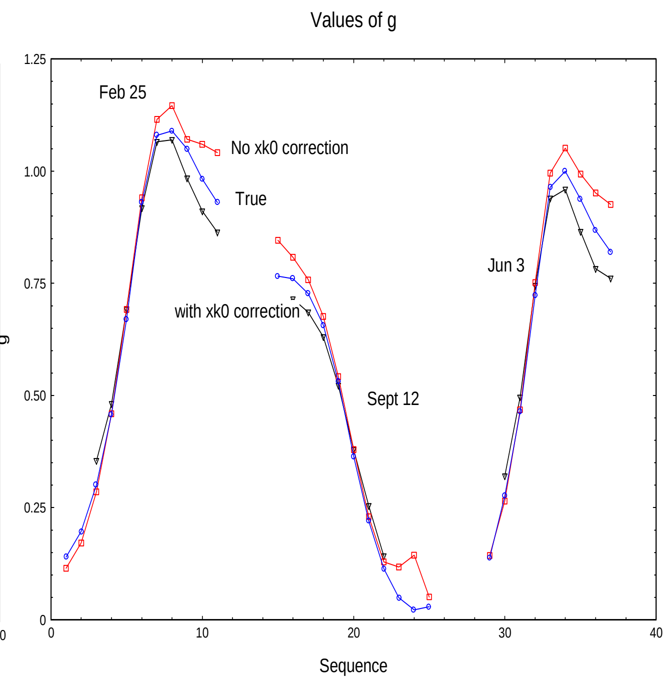


Figure 6

# Methodology Used

The basic method used to find the correct values for tec, g, d and xk0 is an automated iterative search. Since a large number of FFTs are required, it is best to use only the data that include the infinite frequency arrival time and the dispersed signal. In this way tec, g and d values for an event with a tec <~ 2 can be found in 2 – 3 minutes on a 1.3GHz PC. Events with a tec ~ 4 will take twice as long and events with a tec ~ 8 will take twice as long again. An adequately accurate value for xk0 can usually be found in two or three more passes.

To run the algorithm;

Initial guesses for d and xk0 are made. It is not necessary that these guesses be very good, however the processing time is reduced if better guesses are available. Then g is assumed to be 0.0 and correlation values for a large range of tec values are derived. There will be two maximums found, corresponding to the X and O signals. The first estimate for tec is the average of the two maximum values and the estimate for g is a linear function of the difference of these two values.

Using this first estimate for g, the geometric mean of the correlation value of the O and X signals for a number of tec values near the estimated value are generated to find a maximum. Then g is changed slightly and the process repeated until the value of g is found which results in the largest possible correlation value of the O and X geometric mean.

Using these values of tec, g and d the infinite frequency arrival time of the high band signal is found. If this value is not the same as the value found for the low band signal, then a new value for d is tried and the whole process is repeated. Iterations continue until the value of d is found which results in the same infinite frequency arrival time for both the low and high band signals.

Finally, using the derived values of g, d, and the estimate for xk0, the values of tec that result in the largest correlation for the O and X signals are found. If these values of tec are not the same then new values of xk0 are tried until the tec values found are the same.

It should be noted that this method relies on being able to find the O and X maximums when g is set to 0.0. While this is true most of the time, often when the value of g is small and the ratio of the X signal to the O signal is also small, the method fails. Improvements to the algorithm to account for this problem are being pursued.

# Results

Using the methodology described above, a number of LAPP signals were analyzed. The resulting correlation versus time plot for a typical event is shown in Figures 7 and 8. The broader high band signal is a result of the overlapping of the O and X signals. While the LAPP pulse width is less than 20 ns, the ~50 ns width seen is defined by the 20 MHz TATR bandwidth.

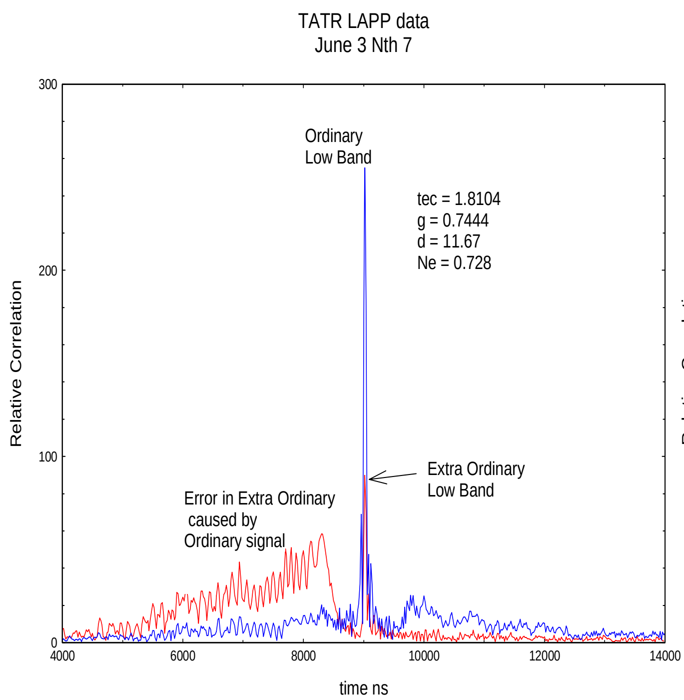


Figure 7

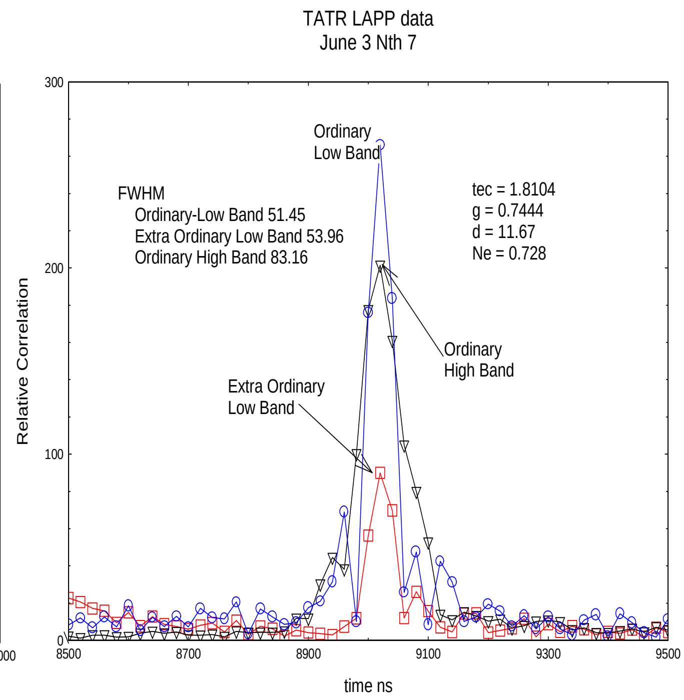


Figure 8

The pulse widths of the 59 events from a total of six passes are plotted as a function of tec and zenith angle in Figures 9 and 10. It is obvious that the widths observed increase as a function of tec and/or zenith angle. It is believed that this increase is caused by:

1) Difficulty in providing good results for small values of g,
2) Decrease in s/n ratio at large zenith angles and/or
3) Failure of the triangular slab model to accurately describe the ionosphere at large zenith angles.

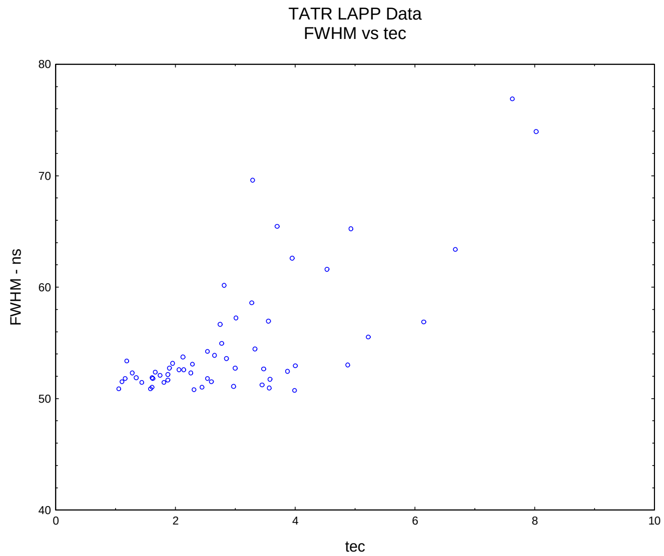


Figure 9

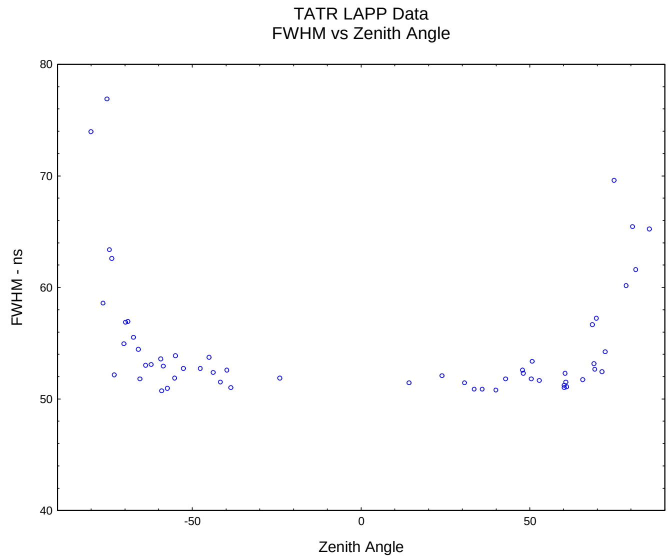


Figure 10

The ratio of the O correlation amplitude to the X correlation amplitude for a series of LAPP passes is plotted in Figure 11. The reason for the ratio being greater than one most (but not all) of the time is unknown. Also the ratio seemed to be essentially constant (at ~ 1.75) on some of the passes but increases with decreasing latitude at other times. These results deserve further study.

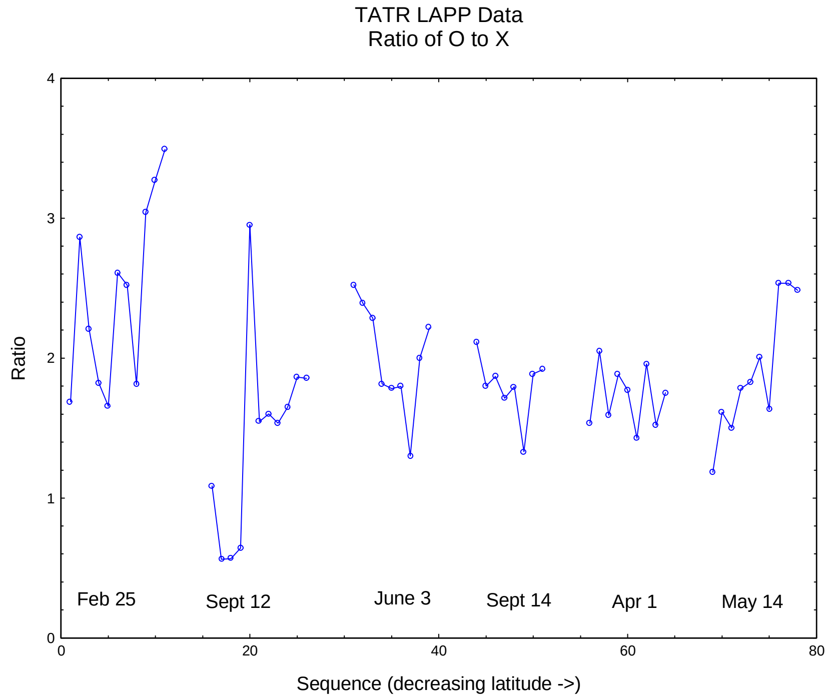


Figure 11

The ionosphere parameters for the 59 LAPP events were derived using the methodology described above. The zenith angle, slant tec and the longitudinal component of the gyro frequency (g) for one of the passes are plotted versus FORTE latitude in Figure 12. The values of vertical tec, slab thickness, peak electron density and the quartic constant d for the same pass are plotted as Figure 13. An indication of the quality of these data can be estimated from:

1) All the parameters should vary smoothly across a pass. Thus any scatter from a smooth curve is an indication of errors in the process. The major causes of such errors are small values of g, low s/n ratios and very large zenith angles. The values of slab thickness and peak electron density, being derived from the small $1/f^4$ term are most subject to these errors.
2) There should be no dependence on zenith angle for the parameters vertical tec, slab thickness or peak electron density. Thus any observed variation indicates an inherent error in the methodology. It seems that this kind of error is minimal for zenith angles less than about 75 degrees.

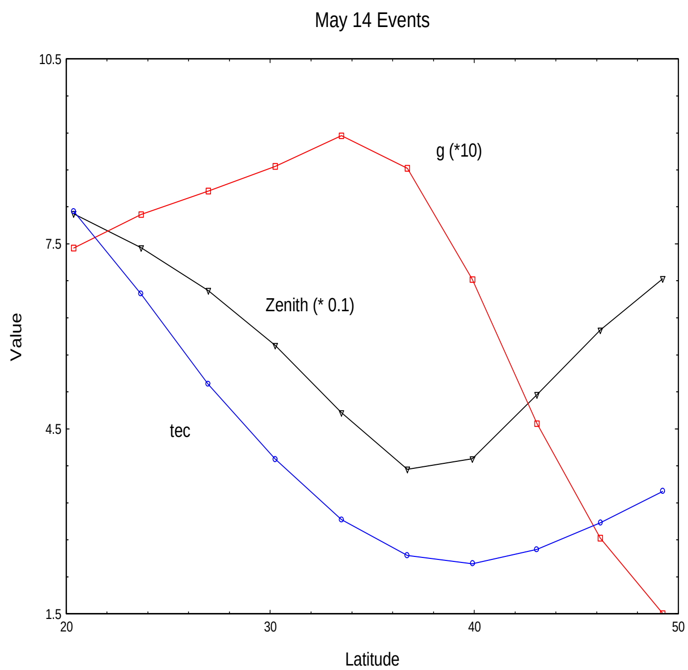

Figure 12

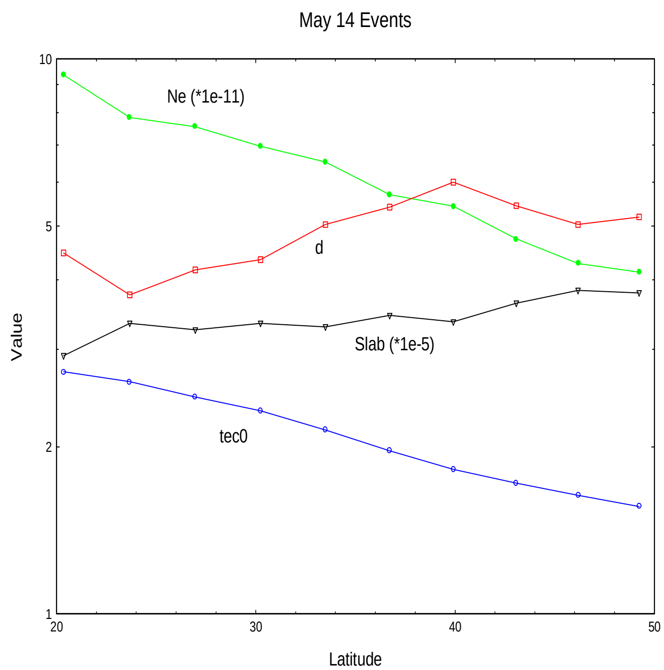

Figure 13

It has been observed (Weidman and Krider, 1980,Levine et al., 1989, Jacobson, 2000 and Jacobson et al., 2001) that CG- strokes over water produce very narrow signals with pulse widths on the order of 60 – 100 ns. With the ability to resolve pulse widths as small as 50 ns from TATR data using this methodology, it should be possible to measure the pulse widths of these events from a satellite. The signals reported by Jacobson (Jacobson, 2000 and Jacobson et al.., 2001) were examined using the methodology and the results are plotted in Figures 14 – 17. It appears that the peak correlation routine introduces artifacts as skirts on the correlation signal, but a pulse width estimate of less than 100 ns but greater than 50 ns is reasonable.

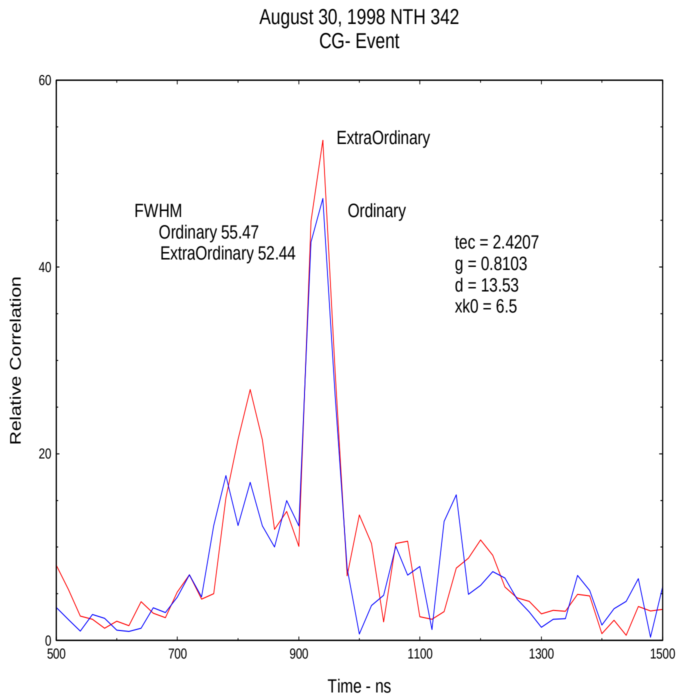


Figure 14

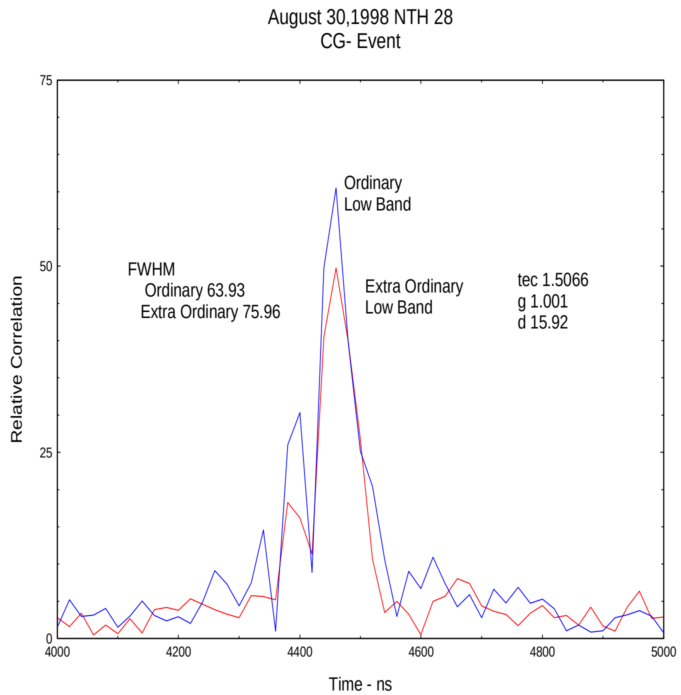


Figure 15

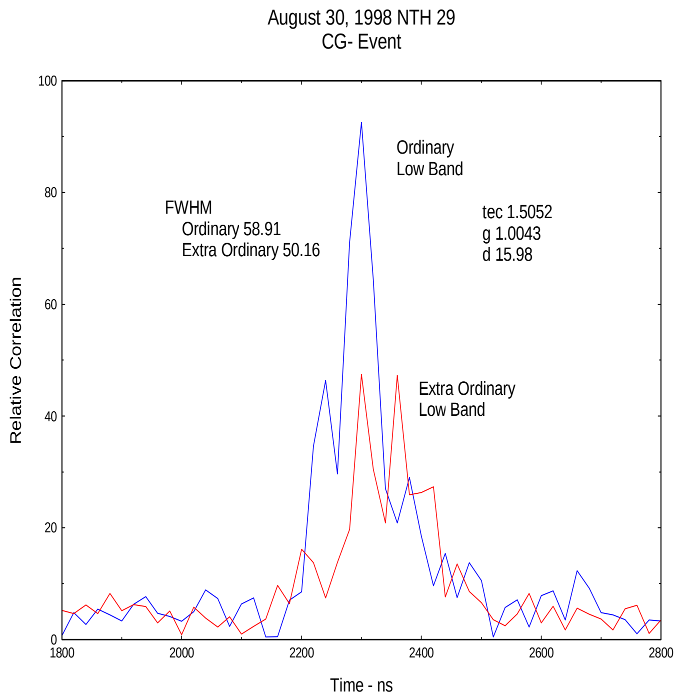


Figure 16

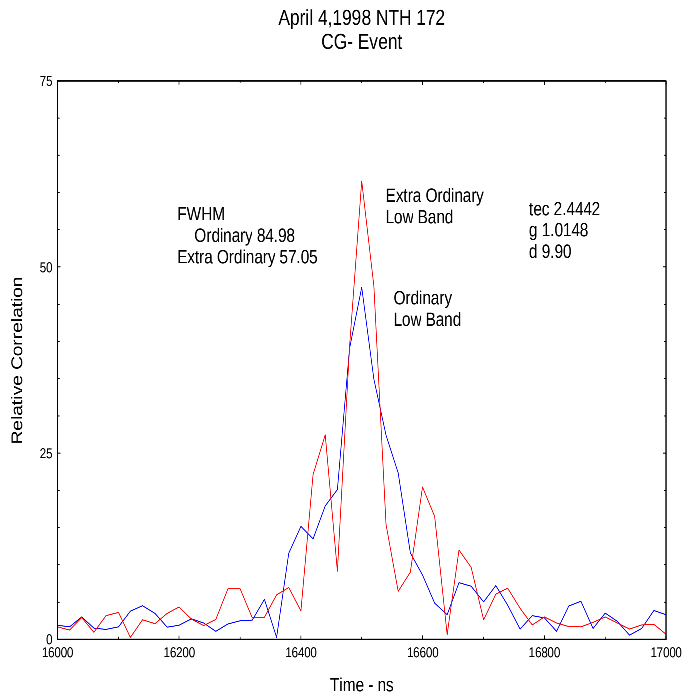


Figure 17

The same methodology has been used to look at the 300 msps HUMR data from the FORTE satellite. The resulting correlation signal from a typical LAPP test is shown as Figure 18. The narrower pulse width (~25 ns) is a result of the broader bandwidth of the HUMR system. It is planned to look at some CG- events to try to derive more accurate values of CG- event pulse widths.

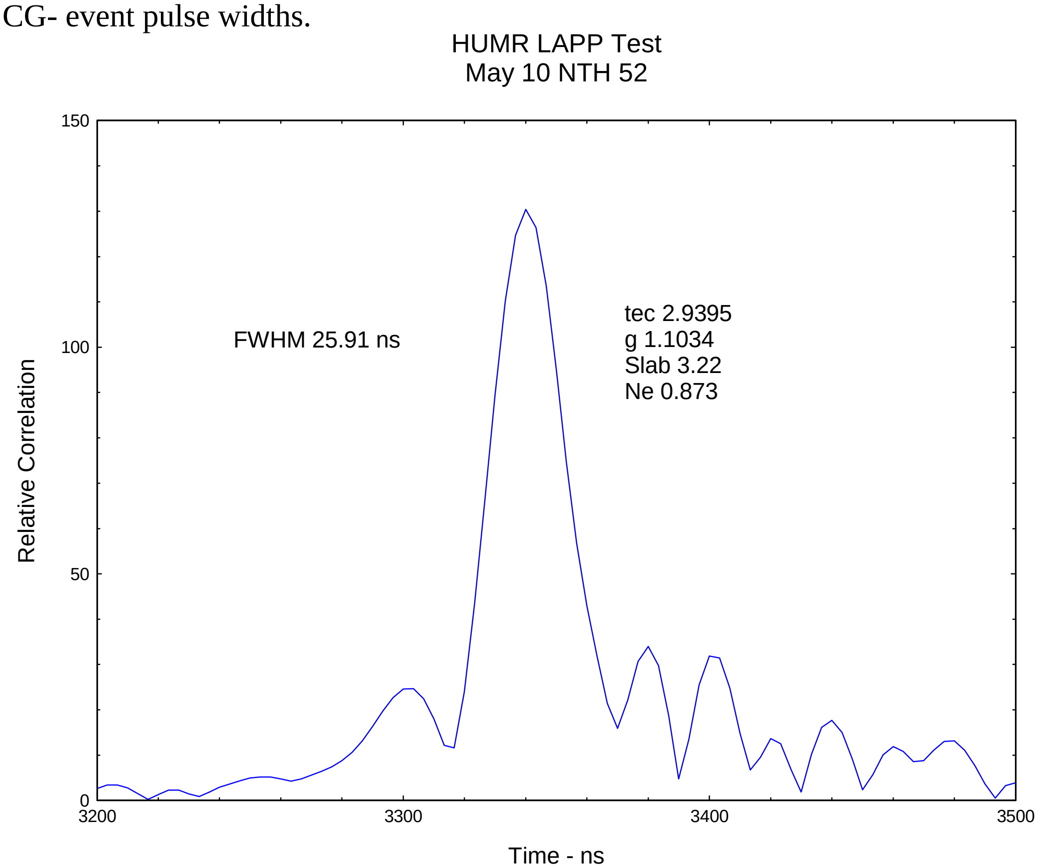


Figure 18

# Acknowledgements

The author wishes to thank Dick Spalding of Sandia and Abe Jacobson and Bob Roussel-Dupre of Los Alamos for their support and many useful suggestions. A special thanks is due Phil Klingner of Los Alamos for his tireless efforts in providing the FORTE data files.

# References


Budden,K.G., The Propagation of radio waves: the theory of radio waves of low power in the ionosphere and magnetosphere, Cambridge University Press, New York, pp. 1-669, 1985

Davies, K. Ionospheric Radio Propagation, National Bureau of Standards Monograph 80, US Government Printing Office, Washington DC, pp 1-470, 1965

Jacobson, A.R. , S. O. Knox, R. Franz, and D. C. Enemark, FORTE observations of lightning radio-frequency signatures: Capabilities and basic results, Radio Sci., 34(2) 337-354, 1999.

Jacobson, A.R. , Flash level relationship of FORTE VHF signals and strokes detected by NLDN, Los Alamos National Laboratory report number LA-UR-00-5976, 2000

Jacobson, A.R. and X.-M .Shao, FORTE satellite observations of very narrow radiofrequency pulses associated with the initiation of negative cloud-to-ground lightning strokes, Los Alamos report LA-UR-01-01-6320, 2001

LeVine, D.M. , J. C. Willett, and J. C. Bailey, Comparison of fast electric field changes from subsequent return strokes of natural and triggered lightning, J. Geophys. Res., 94 (DII), 13259-13,265, 1989

Massey, R.S., S.O. Knox, R.C. Franz, D.N. Holden and C.T. Rhodes, Measurements of transionospheric radio propagation parameters using the FORTE satellite, Radio Science 33, No. 6, 1739-1753, 1998

Roussel-Dupre, R.A., A. R. Jacobson and L. A. Triplett., Analysis of FORTE Data to Extract Ionospheric Parameters, Radio Science vol 36, number 6, pages 1615-1630 (2001).

Weidman, C.D. , and E. P. Krider, Submicrsecond risetimes in lightning return-stroke fields, Geophys. Res. Lett.7(11), 955-958, 1980


The Los Alamos reports can be found at http://nis-www.lanl.gov/nis-projects/forte_science/